

RACE  
Version: v3.1  
Contact: Martin Turner and Mary McDerby  
Date: April 2009



## **Repository of Access Grid Collaborative Events (RACE) University of Manchester**

### **JISC Final Report**

*Martin Turner and Mary McDerby  
23<sup>rd</sup> April 2009*



## Repository of Access Grid Collaborative Events (RACE) University of Manchester

### JISC Final Report

Project Information			
<b>Project Acronym</b>	RACE		
<b>Project Title</b>	Repository of Access Grid Collaborative Events		
<b>Start Date</b>	1 October 2007	<b>End Date</b>	31 March 2009
<b>Lead Institution</b>	The University of Manchester		
<b>Project Director</b>			
<b>Project Manager &amp; contact details</b>	Martin Turner Research Computing Services, Kilburn Building, The University of Manchester, Manchester M13 9PL Tel. 0161 275 0613 email: martin.turner@manchester.ac.uk		
<b>Partner Institutions</b>	None		
<b>Project Web URL</b>	<a href="http://www.rcs.manchester.ac.uk/research/race">http://www.rcs.manchester.ac.uk/research/race</a>		
<b>Programme Name (and number)</b>	<i>Repositories start-up and enhancement projects</i>		
<b>Programme Manager</b>	Andrew McGregor		

Document Name			
<b>Document Title</b>	<i>Final Report</i>		
<b>Reporting Period</b>	<i>End of Project</i>		
<b>Author(s) &amp; project role</b>	Martin Turner, Project Manager; Mary McDerby, Developer		
<b>Date</b>	20 <sup>th</sup> February 2009	<b>Filename</b>	RACEFinalReport
<b>URL</b>	N/A		
<b>Access</b>	<input checked="" type="checkbox"/> Project and JISC internal		<input type="checkbox"/> General dissemination

Document History		
Version	Date	Comments
2.0	17 <sup>th</sup> March 2009	Review
3.0	31 <sup>st</sup> March 2009	Pre-Final Review 2
3.1	14 <sup>th</sup> April 2009	Corrections and dissemination list



## JISC Final Report – RACE

- Web (JISC and local)
  - [http://www.jisc.ac.uk/whatwedo/programmes/programme\\_rep\\_pres/repositories\\_sue/race.aspx](http://www.jisc.ac.uk/whatwedo/programmes/programme_rep_pres/repositories_sue/race.aspx)
  - <http://www.rcs.manchester.ac.uk/research/race>
- Wiki
  - <http://wiki.rcs.manchester.ac.uk/community/RACE>
- Blog
  - <http://grace.rcs.manchester.ac.uk/race-blog/>

## Table of Contents

Acknowledgements .....	3
Executive Summary .....	4
Background .....	5
Aims and Objectives .....	5
Methodology .....	6
Implementation.....	7
Outputs and Results .....	8
Outcomes .....	9
Dissemination.....	10
Conclusions.....	10
Implications and Sustainability .....	11
Recommendations .....	11
References .....	11

## Acknowledgements

This project was funded by JISC under the Start-up and Enhancement Projects Training (SUETr) programme.

We also acknowledge the e-Dance project, funded by the joint [AHRC+EPSRC+JISC e-Science Initiative](#); CREW project, funded by the JISC under the second phase of the [Virtual Research Environments \(VRE\)](#) programme, and the UK ACM SIGGRAPH Professional Chapter.

- <http://projects.kmi.open.ac.uk/e-dance/>
- <http://www.rcs.manchester.ac.uk/research/collaborativeresearcheventsontheweb> or <http://www.crew-vre.net/>
- [http://wiki.rcs.manchester.ac.uk/community/ACM\\_SIGGRAPH\\_Chapter](http://wiki.rcs.manchester.ac.uk/community/ACM_SIGGRAPH_Chapter) or <http://manchester.siggraph.org/>

## Executive Summary

This project was to develop a digital repository service of collaborative research events such as seminars, conferences, workshops, performances, and training and teaching events. The repository service aims were to archive such sessions and allow the content to be searched, browsed through and retrieved via existing and established repository search services.

The main features were:

- Set up a repository service for Access Grid (next generation of video conferencing) collaborative events.
- Define a metadata element set appropriate for AG data that allows for efficient discovery.
- Produce a set of recommendations regarding potential modifications to the AG recording software being developed within the context of the JISC-funded CREW project.
- Achieve successful and efficient integration with the Intute Repository Search Service via CREW

Two pioneer user groups (the UK ACM SIGGRAPH University of Manchester Professional Chapter and a group of performance artists, specialising in contemporary choreography that have been brought together by the AHRC funded e-Science project e-Dance) were engaged throughout the lifetime of the project so that their needs and work practices played an important factor in the development plan. To that end, they also played an important role as evaluators and their recommendations were regularly reviewed by a Steering Committee, headed by the Project Manager and including representatives of the development and the pioneer user groups.

The content within the repository has been made centrally available to the two pioneer groups, both of which have been enthusiastic exploiters of early prototype Access Grid recording systems. The much wider end target audience has been the UK AG community, Intute and, from this, all researchers that make use of the Intute Repository Search Service for retrieval of archived research material.

The outcomes of the RACE project will continue to be used by the e-Dance project as a repository for performances and rehearsals. Another outcome is metadata standards for video, which is also currently a project proposal under review.

## 1 Background

This project was to develop a digital repository service of collaborative research events such as seminars, conferences, workshops, performances, and training and teaching events. The repository service archives such sessions and allows the content to be searched, browsed through and retrieved via existing and established repository search services.

The project also aimed to build on the very promising software within the context of 'Collaborative Research Events on the Web (CREW)', a project funded by JISC within the second phase of the Virtual Research Environments (VRE) programme that is capable of recording and annotating Access Grid (AG) events. CREW will be integrated with Intute to allow for the repository to be searched via the Intute Repository Search Service.

The content within the repository has been made centrally available to two pioneer groups, both of which have been enthusiastic exploiters of early prototype Access Grid recording systems; the UK ACM SIGGRAPH Professional Chapter and a group of performance artists, specialising in contemporary choreography that have been brought together by the AHRC funded e-Science project e-Dance. The much wider end target audience has been the UK AG community, Intute and, from this, all researchers that make use of the Intute Repository Search Service for retrieval of archived research material.

Such a repository was also important as the AHRC Data Service Repository ceased to be fully funded around the same time as the startup of this project. The RACE repository allowed the e-Dance project to store footage of rehearsals, performances and other such important information as background details on the performances, access grid environment etc.

Experience from the team had overlapping skills and brought relevant *a priori* previous work to the project:

- Project management of the CREW and e-Dance projects
- JISC funded National Video Facility – which included the digitization of video footage as well as real-time streaming. This was a national service.
- Access Grid Support Centre (AGSC) – working with collaborative research events within a service environment.

## 2 Aims and Objectives

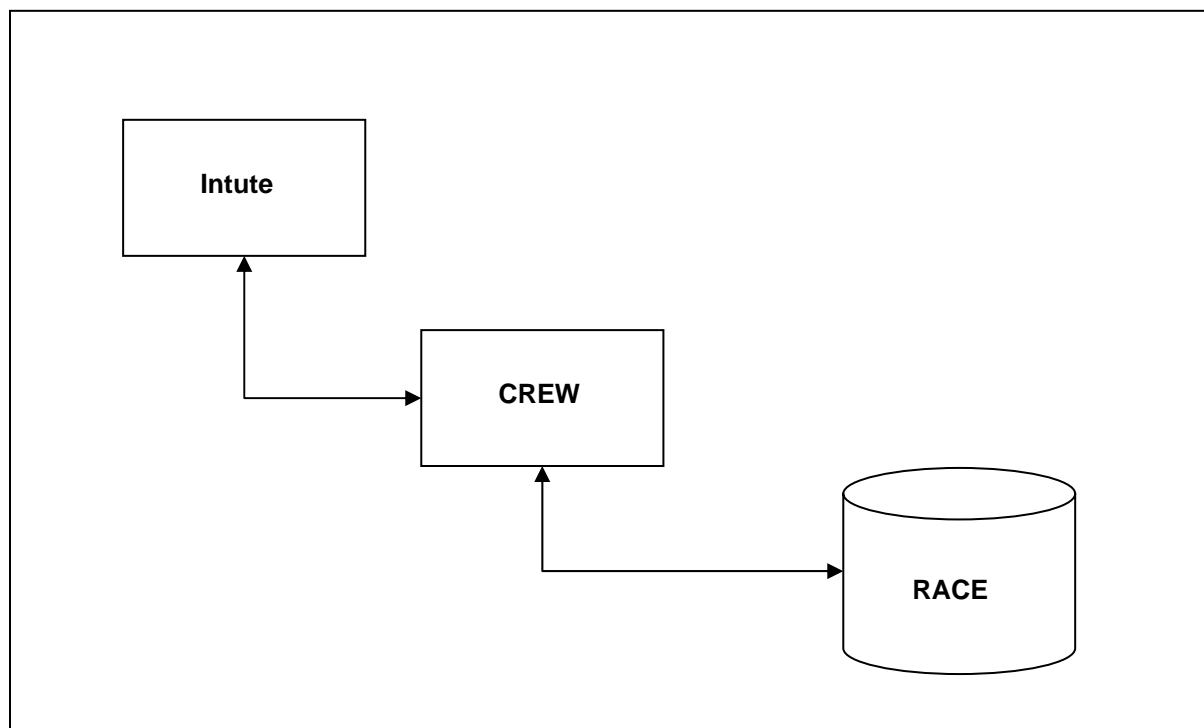
The aim of this project was to implement a repository service for AG collaborative events such as conferences, performances, workshops and seminars; discoverable and searchable by existing repository search tools.

The specific objectives we intended to achieve were:

- Set up a repository service for AG collaborative events.
- Define a metadata element set appropriate for AG data that allows for efficient discovery.
- Produce a set of recommendations regarding potential modifications to the AG recording software being developed within the context of the JISC-funded CREW project.
- Achieve successful and efficient integration with the Intute Repository Search Service via CREW

An investigation was undertaken into the available repository software with regards to the needs of Access Grid recordings. ePrints, Fedora Commons and DSpace were explored. DSpace was found to be the most applicable to our needs – intuitive, customizable i.e., front end, metadata etc, API frontend – allowing us to communicate directly with CREW/upload via a simple interface.

The initial structure of the project was that RACE would store all AG content (video, audio etc) – CREW would be the frontend to RACE – populating the repository, and Intute would harvest metadata regarding the AG content from CREW (Figure 1).



**Figure 1: Initial Project Structure**

The metadata element was defined – first as a general Access Grid dataset for CREW recordings, then as a more specific set was defined for the e-Dance users. Unfortunately, it was found that the use did not quite fit the metadata standard. But, DSpace allows you to override such shortcomings. Metadata with regards to digital recording is a long way off – and beyond the scope of this project to start defining in full.

Recommendations regarding potential modifications to AG recording in the context of CREW took the form of a user based involvement. RACE attended CREW user days and fed into the user requirements gathering.

The major change to the aims and objectives was that the RACE repository could not be used with CREW due to the project being driven down the semantic web route. Therefore, the development of CREW went in a different direction – the CREW recording interface changed and the semantic web metadata did not fit with the metadata functionalities within RACE. However, the recommendations we suggested on metadata in CREW were adhered to. Therefore, with this major change our user communities became e-Dance and ACM SIGGRAPH.

### **3 Methodology**

The development team evaluated different digital archival systems, followed by setting up a repository service and integrating it with the AG recording software. The main methodology was that of construction, test, evaluation and re-construction. This sequence was visited twice over the lifetime of the project to improve the repository service.

User workshops, usability sessions and meetings with the pioneer user groups were the main means for testing. To allow for this, two pioneer user groups were actively engaged throughout the lifetime of the project, as opposed to having a mere responsive role, so that their needs and work practices constitute an important factor in the development plan. The following communication tools were also created and setup:

- Web (JISC and local),

- [http://www.jisc.ac.uk/whatwedo/programmes/programme\\_rep\\_pres/repositories\\_sue/race.aspx](http://www.jisc.ac.uk/whatwedo/programmes/programme_rep_pres/repositories_sue/race.aspx)
  - <http://www.rcs.manchester.ac.uk/research/race>
- Wiki
  - <http://wiki.rcs.manchester.ac.uk/community/RACE>
- Blog
  - <http://grace.rcs.manchester.ac.uk/race-blog/>

### **3.1 Important issues**

The following are the important issues for this project:

- Standards-based. All software used and produced through the lifetime of the project is based on common and, as far as possible, non-proprietary standards.
- Usability. This was a critical factor in order to create a service that will be used by real-life users with their associated needs and practices. This also aided in drawing plans for sustainability beyond the lifetime of the project. The user has had an important role in the project methodology for these reasons.
- Interoperability. The repository was designed to allow for efficient interoperability with related tools such as repository search services.

### **3.2 Scope and boundaries of work**

The scope and boundaries of RACE was to implement a repository service of AG collaborative events and integrate with repository search services such as the Intute Repository Search via CREW. A standalone frontend was also developed to enable ease of use of uploading materials to the repository by the pioneer users. This involved investigation into the user requirements and needs.

### **3.3 Critical success factors**

The key factor for success was targeted as:

- Usability – is the software/service easy and intuitive?
- Visibility – is the service visible to external repository search tools?
- Acceptance – do people use it?

## **4 Implementation**

To meet the aims and objectives outlined above the following was undertaken in order to set up the repository service.

- An investigation of Digital Archival systems was undertaken. The systems which were evaluated were DSpace, ePrints and Fedora Commons. This involved looking at the usability, ability to customize, whether an API was available, ease of use of the API – for example could it be easily used to upload material via. The systems evaluated were test driven with the pioneer users to determine their usability and whether the software was suitable. Further evaluation was undertaken with the CREW/e-Dance developers who would use the API to upload their materials to the repository. A key issue that was discussed in these meetings was that we did not want to re-invent the wheel by using repository systems such as Fedora Commons.

Further discussion was undertaken with the University of Manchester Institutional Repository Project team – who started work on their project at the same time as RACE did. The University repository team were also looking at DSpace, ePrints and Fedora Commons in order to create a repository to store digital copies of scholarly work. This was quite important as integration with RACE could be a possibility in the future.

DSpace was chosen for the RACE repository due to its intuitiveness, ease of use (from user and system administrator point of view), API, ability to customize metadata schemas (easily),

etc. This decision was undertaken in collaboration with the two pioneer user groups, as well as the CREW and e-Dance developers.

The University of Manchester Institutional Repository will take a hybrid approach incorporating the use of all three repository systems mentioned above. This means that in future RACE could be integrated with that project.

- Dublin Core was taken as the metadata element set due to the fact it is a standard and the broad range of its capabilities. This was the starting point for describing the Access Grid materials – which can be quite complex. There is currently no standard for video metadata – as research continues in this field (<http://microformats.org/wiki/video-metadata-model>). However, with regards to our pioneer users the metadata was even more complex – combining video metadata with simple Dublin Core, and going one step further into including a form of “spatial awareness” type metadata – Access Grid setup information i.e., window positions on projection walls, camera positions, camera positions in relation to the window positions etc, etc. DSpace allowed us to extend the Dublin Core metadata standard (within the standard’s guidelines) to include this information.
- There were some issues with the DSpace API and the customized metadata standard. First of all the extended metadata element set was undertaken as administrator within the DSpace interface. However, documentation fails to communicate the number of configuration files that must be edited in order for the API to know about these new metadata elements. Our blog fully reports on these areas (<http://grace.rcs.manchester.ac.uk/race-blog/>). Also full instructions are listed on how to do it yourself!

## 5 Outputs and Results

Key deliverables including software and services, reports etc are listed below.

### 5.1 Software and Services

- Metadata set for e-Dance and CREW
- Hardware setup for AG repository service with large amount of disk space for AG recordings.
- AG repository service using DSpace tested and put into service. The service is maintained, and housed in a 24x7 monitored machine room.
- Standalone frontend for e-Dance users to upload their video material. This was created using the DSpace API.
- Communities setup on AG repository service for pioneer user groups.
- e-Dance are currently preparing 25 gigabytes of material to upload onto the RACE repository.
- Support material for the above outputs (documentation, presentations, etc.)

### 5.2 Reports

1. Recommendation of a digital archiving system for Access Grid (AG) recordings.
2. Recommendations of a metadata element set for Access Grid (AG) recordings.
3. Service requirements, based on user needs and practices.
4. Recommendations suggesting any required modifications to the e-Dance frontend.
5. Repository service usability review.

All reports are kept on the RACE website:

<http://www.rcs.manchester.ac.uk/research/race>

Items 3, 4 and 5 are combined in one report for e-Dance.

### 5.3 Other

- Web (JISC and local), wiki and Blog created and setup
  - [http://www.jisc.ac.uk/whatwedo/programmes/programme\\_rep\\_pres/repositories\\_sue/race.aspx](http://www.jisc.ac.uk/whatwedo/programmes/programme_rep_pres/repositories_sue/race.aspx)
  - <http://www.rcs.manchester.ac.uk/research/race>
  - <http://wiki.rcs.manchester.ac.uk/community/RACE>
  - <http://grace.rcs.manchester.ac.uk/race-blog/>
- Online usability questionnaire:
  - <http://www.rcs.manchester.ac.uk/research/race/usereval>

## 6 Outcomes

The RACE repository will improve the value of AG collaborative sessions as these will have the potential to become lasting research materials including valuable metadata information. Students, academics and researchers will benefit, as this new rich form of research material will be available to support their activities. The AG community will also grow and advance as a result.

With the funds drastically reducing for the Arts and Humanities Data Service (AHDS), this new repository service fulfils some of the immediate needs of users such as the e-Dance community, who record hours of material whilst rehearsing performances. Such information would be easily lost if not stored/catalogued. The metadata schema created for the e-Dance users reflected more of the wealth of information which could be lost – describing camera and window positions etc in the context of spatial awareness and use of the Access Grid. Our belief throughout the creation of the metadata schema was to give the user as much information as possible for them to search the repository with – this was thought of as a necessity due to the complexity of the material.

*Concerns raised by user questions specifically regarding future use of RACE:*

- IPR and confidentiality – we are dealing with recording multiple people in multiple sites to be replayed at multiple other sites.  
*This is being addressed through a ja.net initiative launched through the AGSC (Access Grid Support Centre). Current recordings are classified as tests and still being recorded.*
- Reliability of service and guaranteed quality – mainly due to recording at a remote server. This has failed a few times on test cases: as we may lose data.  
*This issue has been raised with the CREW developers and with the AGSC team. A process of stable versions is proposed to remove excessive use of experimental tools.*
- Buy-in and use of metadata standards by the users. Great concern about users miss-typing or not typing standard forms.  
*Early days for this issue as minimal annotation is currently being carried out except for some of the core metadata.*

*Developer Issues:*

A recent issue from developers is the future definition of metadata. New recording software is now being made web 2.0 compliant. This can include features from the Semantic Web philosophy to be incorporated. There is a lack of definition for annotation which now needs to consider comments and timing markings, as well as incorporate the formal metadata. It was because of these problems that CREW decided to use their own approach to a repository system (an inhouse developed piece of software – created in Bristol).

*Hardware issues:*

Hardware servers from the British Library were due to arrive and be built and installed in May 2008. However, this did not happen. We did chase – and after several conversations we were told that the British Library had employed a new system administrator who wanted to keep the equipment. So we went into our contingency plan and bought several SATA drives to add to one of our existing servers in July 2008.

### *IPR Issues*

A report is being undertaken by one of our collaborative projects – CREW on IPR. This will cover any issues that would be encountered by the RACE project.

## **6.1 Dissemination**

- Attended DSpace Technical Briefing Day, Bristol 12th November 2007
- Attended SUE Programme Meeting 27<sup>th</sup>-28<sup>th</sup> November 2007
- Attended JISC Project Management Meeting 4<sup>th</sup>-5<sup>th</sup> December 2007
- User info-gathering session: Intute users (Manchester) - initial meeting 23rd January 2008
- Internal University of Manchester RACE demonstration day - 5th February 2008: showcase test repository servers to potential developers.
- User information gathering session: Partner requirements from JISC VRE (Virtual Research Environment) Phase 2 CREW project (Manchester plus email to Bristol).
  - Including user evaluation day for scientific users 4<sup>th</sup> October 2007.
  - CREW structure meeting 4<sup>th</sup> November 2007 and CREW architecture day - cross-networking opportunity.
- Attended Forum on Research Data Management - held at the Chancellors at the University of Manchester. 19<sup>th</sup>-20<sup>th</sup> March 2008
- Contact formed with University of Manchester Institutional repository project
- *Work package 2* – feedback sessions were held with e-Dance and ACM Siggraph Manchester Chapter. (This is an iterative process throughout the whole of the project).
- Meetings have been held on:
  - “Feedback session” at Bedford e-Dance 10th-11th April 08
  - CREW users meta-data day - 28th April also “feedback session”
  - CREW developer meetings – weekly
- Access Grid Retreat, June 2008, Joint paper with CREW project.
- e-Science All Hands Meeting (Edinburgh September 08) – Poster and leaflet with CREW project.
- 3rd November 2008 SurfNET SURF Academy on online collaboration
- 6th November 2008 JISC sponsored event: Advanced Tools and Technologies for Collaborative Research
- 19th February 2009 – RACE Workshop (attendees included pioneer users, CREW developers, PAG developers, Access Grid Support Centre staff, University of Manchester – Streaming Manchester staff etc).
- 24<sup>th</sup> March 2009 – Included element of CREW (VRE2) demonstration at JISC Conference
- 23<sup>rd</sup> April 2009 – University of Manchester Institutional Repository Launch. Presentation on RACE.
- 6<sup>th</sup> and 7<sup>th</sup> May 2009 – JISC Repositories and Preservation Programme Meeting, Aston Business School, Birmingham.
- 9<sup>th</sup> June 2009 – JISC VRE2 final Programme Meeting, London Links with VRE2-CREW

## **7 Conclusions**

The result of the RACE project will continue to be used by the e-Dance project as a repository for its performances and rehearsals. The success of this project has led to its results being continued to be used in a proposed follow-on project to further the e-Dance work.

The issues raised such as metadata standards and IPR need to be further progressed before such a repository could become a mainstream player in the Access Grid community. However, it is felt that

such a repository and metadata schema is urgently needed in order to allow Access Grid to advance further and widen its community.

With the launch of The University of Manchester Institutional Repository (23<sup>rd</sup> April 2009), RACE has the possibility of being carried forward as a University wide service for Research Data i.e., group experimental findings, large data etc. The University repository intends only to hold scholarly articles. It is thought that RACE can aid with the increasing problem of shared storage resources for projects; advanced collaborative working within a project etc.

## 8 Implications and Sustainability

The University of Manchester is committed to maintaining the medium term use of the local system and nationally through related networks.

### *Future Research/Development Projects*

- The result of the RACE project will continue to be used by the e-Dance project as a repository for its performances and rehearsals.
- Continued use by the e-Dance project will provide extensions to the comprehensive meta-data creation and capture will be undertaken, improving the repository interface.
- The success of this project has led to its results being continued to be used in a proposed follow-on project to further the e-Dance work.
- With the launch of The University of Manchester Institutional Repository, RACE has the possibility of being carried forward as a University wide service for Research Data i.e., group experimental findings, large data etc. The University repository intends only to hold scholarly articles. It is thought that RACE can aid with the increasing problem of shared storage resources for projects; advanced collaborative working within a project etc.
- Future work is needed on Metadata Standards for Access Grid Environments. A wider consultation and longer evaluation of uses for these new metadata terms need to be considered as future work.

**Mary McDerby will continue to administer and maintain the RACE project in collaboration with e-Dance staff.**

## 9 Recommendations

DSpace although a stable and usable repository system, needs advanced user documentation outlining API communication – especially with regards to customized metadata schemas. We have uploaded a few simple guidelines at <http://grace.rcs.manchester.ac.uk/race-blog/>

Metadata standards for video are a current research topic. However, metadata with regards to Access Grid environments needs more information and should be a thought for the future.

## 10 References

All projects and locations are referenced within the context of this document.